The Silenced Text: Field Experiments on Gendered Experiences of Political Participation

Alan N. Yan[1*] and Rachel Bernhard[2]

[1]Department of Politics, Princeton University,
Fisher Hall, Princeton, NJ 08544, USA
[2] Department of Political Science, University of California, Davis,
469 Kerr Hall, One Shield Avenue, Davis, CA 95616, USA

*Corresponding author's email: anyan@princeton.edu.

Campaign volunteers assigned female names receive more offensive and silencing messages than ambiguous or male names.

Who gets to "speak up" in politics? Whose voices are silenced? While much research examines incivility in public discourse between and towards elites, little examines how incivility might shape the everyday experiences of politics for men and women in the U.S. today. We conducted two field experiments in which we randomized the names campaign volunteers used to text voters reminders to participate in a protest and call their representatives. We find that female-named volunteers receive more offensive, discouraging, and silencing responses than male-named or ambiguously-named volunteers. We suggest these findings help make sense of prior research finding women are less likely to participate in politics than men, and conclude by discussing the implications for political activism and civic debate.

Introduction

In the wake of the 2016 U.S. presidential campaign, both journalists and academics have expressed concern over the growing incivility in public discourse. Though substantial research documents the consequences of incivility between elites to voter opinion (*1*), we are only beginning to understand how such interactions impact public trust in democracy (*2, 3*). We know little about incivility between citizens in political deliberation, and nothing at all about how these attitudes and behaviors play out in a new and immensely popular form of political participation: text canvassing (*4*).

This lacuna matters because women may be more likely than men to face uncivil behaviors when they enter the public sphere (*5*). Documented behaviors run the gamut from the irritating, like interruption and silencing (*6, 7*) to the dangerous, such as sexual harassment and violence (*8*). Research on gendered harassment typically focuses either on its prevalence in the home and the workplace—excluding politics—or, if on politics, on the treatment of women leaders. In neither case do we learn what everyday activism is like in the U.S. today. We therefore ask a simple question: are women more likely than men to receive uncivil messages when they participate in politics?

To examine the experiences of men and women participating in political exchanges, we conducted two field experiments in which we randomized the names used by volunteers during a texting ("Short Message Service," or SMS) campaign meant to encourage a liberal organization's supporters to attend rallies or call their representatives. Volunteers used a software program to rapidly text voters pre-written messages, enabling randomized assignment of names not just across volunteers but within volunteers across respondents, removing the possibility of systematic differences between volunteers. For each supporter, a volunteer was randomly assigned to use a male name, a female name, an ambiguously-gendered name, or no name. This

manipulation allows us to identify perceived gender as one cause of men and women's different experiences in politics (*9*).

In both studies, female-named volunteers receive more offensive, discouraging, and silencing responses than male-named or ambiguously-named volunteers. These findings are concerning since many barriers to participation already exist, especially for historically disadvantaged groups like women (*10*), ethnic minorities (*11, 12*), and the poor (*13*). Previous research has been unable to explain why women are more likely than men to vote but less likely to participate in other political activities after holding constant important factors like education and income (*14*). Our findings help make sense of women's reluctance to participate in peer-oriented political activities like canvassing, and in turn raise new questions about women's equality in everyday political life in the U.S., during the 2020 elections and beyond.

Results

We conducted two randomized control trials (RCTs) in 2018 evaluating whether female-named volunteers receive more harassing SMS responses than male-named volunteers. We partnered with political action committee NextGen America (NGA) to obtain samples of individuals who had previously interacted with the organization: 60,356 in the first trial, 75,231 in the second. In the first RCT, we used participating volunteers' own ratings (self-reports) of how offensive the messages were. In the second RCT, two non-participating volunteers rated how offensive and discouraging they found the messages received by those participating. Finally, we report how many volunteers were silenced—asked to stop contacting the respondent—across both studies.

In the main paper, we provide figures based on ordinary least squares (OLS) regressions. These regressions estimate the treatment effect of assigning someone at phone number i a particular name condition on the rate of offensive or discouraging replies. We employ the "ambiguous-gender" condition as the control condition (*15*), and report whether female and

Table 1: Sample Responses by Coding Category

Offensive	"No. Your fat." "Fuck off" "Piss off" "For the love of god I'm not fucking Edwin you have the rong number and I'm a 9 year old bitch"
Encouraging/Discouraging	"Thanks for your work Jessica!" "All politicians are scumbags - could care less" "That seems like a fruitless Effort" "Having said that, I appreciate your volunteer work and dedication to eliminating government corruption and corporate pollution! Well done, Taylor!"
Silencing	"Take my number off of your volunteering lists or I will file for harassment against the company you are volunteering for. I no longer want to receive any messages from you guys. Thanks." "STOP TEXTING" "No thank you. Please, quit texting" "Lose my number lady"

Codes are non-exclusive: a text may be coded as any combination of these categories.

male voters reacted differently (*16*). In the appendices, we also estimate the average treatment effect for treated phone numbers (those that received a text, rather than all numbers) using instrumental variable regressions; these show larger estimates (*17*).

In both studies, we find that when volunteers are assigned a female name, they are more likely to receive offensive, discouraging, and silencing messages, reducing their ability to participate in the deliberation on which democracy rests (*18, 19*). In (Tab. 1), we display examples of each type of message.

Offensive and Discouraging Replies

In both studies we find that volunteers are more likely to receive offensive replies when assigned to use female names than when assigned to use male or ambiguous names. We do not find a

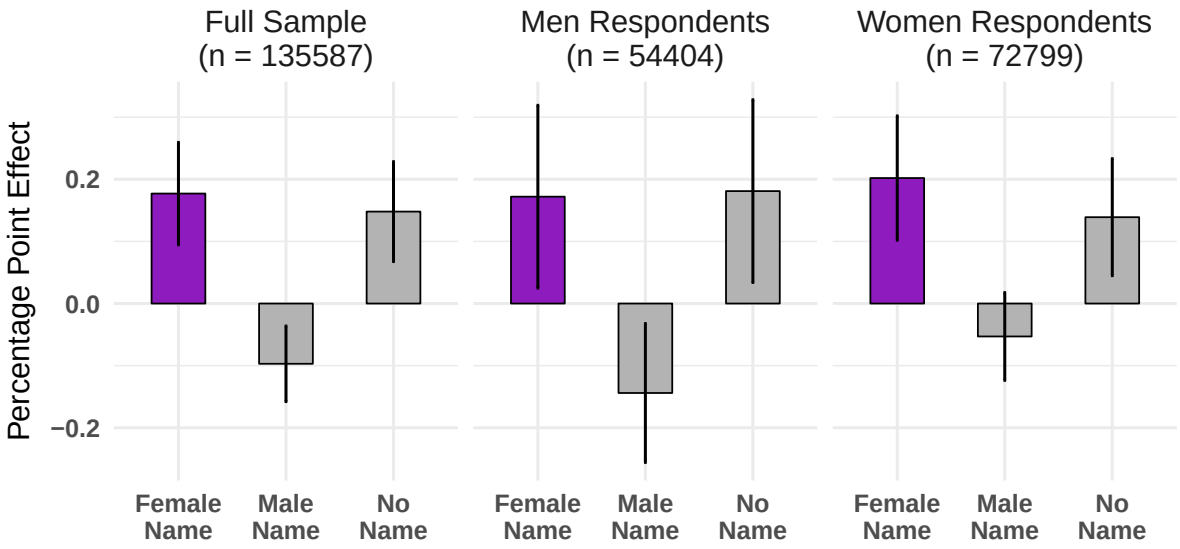

Figure 1: Mean Offensiveness. Figure shows the average treatment effect by treatment condition with 95% confidence intervals estimated using OLS. The comparison category is the ambiguous name condition. Data is from both studies. Table S8 displays the regression estimates depicted above.

significant difference in the ratings for offensiveness between the two studies; we therefore depict the results for both studies below. Tables S9-S11 show the results hold for each study individually and with the addition of volunteer fixed effects.

(Fig. 1) presents the OLS regression results overall and by gender of the respondent for both studies. The left-hand panel reports results by treatment condition for the full sample. Volunteers assigned to use female names are 0.177 percentage points more likely than ambiguously-named volunteers to receive offensive messages during a campaign (two-tailed $p = .000$). Male-named volunteers were 0.097 percentage points less likely to receive offensive messages than the ambiguously-named ($p = .012$). Unnamed individuals were also more likely (.148 percentage points) to receive offensive messages ($p = .000$). The middle and right-hand panels report results for men and women separately. We see no evidence of heterogeneous treatment effects. On average, for every 1,000 messages a female-named volunteer sends, she receives 1.77 more

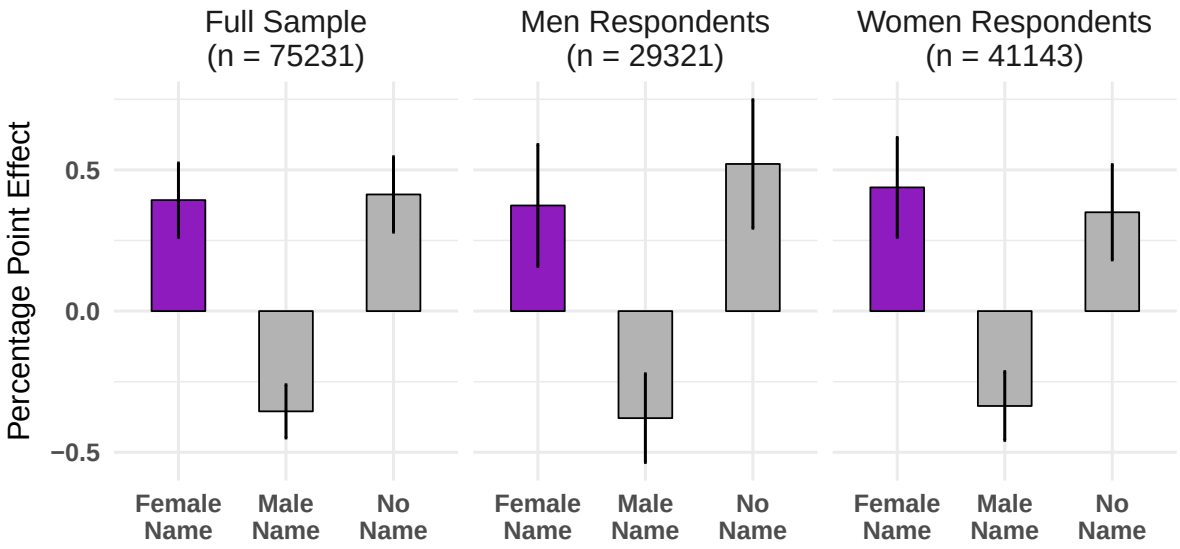

Figure 2: Mean Discouragingness, As Rated by Other Volunteers. Figure shows the average treatment effect by treatment condition with 95% confidence intervals estimated using OLS. The comparison category is the ambiguous name condition. Data is from Study 2 only. Table S12 displays the regression estimates depicted above.

offensive messages than an ambiguously-named texter, and 3.17 more offensive messages than a male-named texter ($p = .000$). In our sample, the average volunteer sent nearly 1,300 texts, suggesting that even a casual volunteer will endure a number of offensive messages.

Study 2 provides evidence that female-named volunteers receive more discouraging replies than the other named conditions (Fig. 2). In the left-hand panel, we show that female-named $(p < .001)$ and unnamed volunteers $(p < .001)$ were significantly more likely to receive discouraging replies than an ambiguously-named volunteer, while male-named volunteers were significantly less likely to receive discouraging replies $(p < .001)$. In the middle- and right-hand panels, we again see no difference in treatment effect by respondent gender. Overall, female-named volunteers can expect to receive 3.93 more discouraging messages per 1,000 messages than the ambiguously-named, and 7.48 more discouraging messages than male-named volunteers.

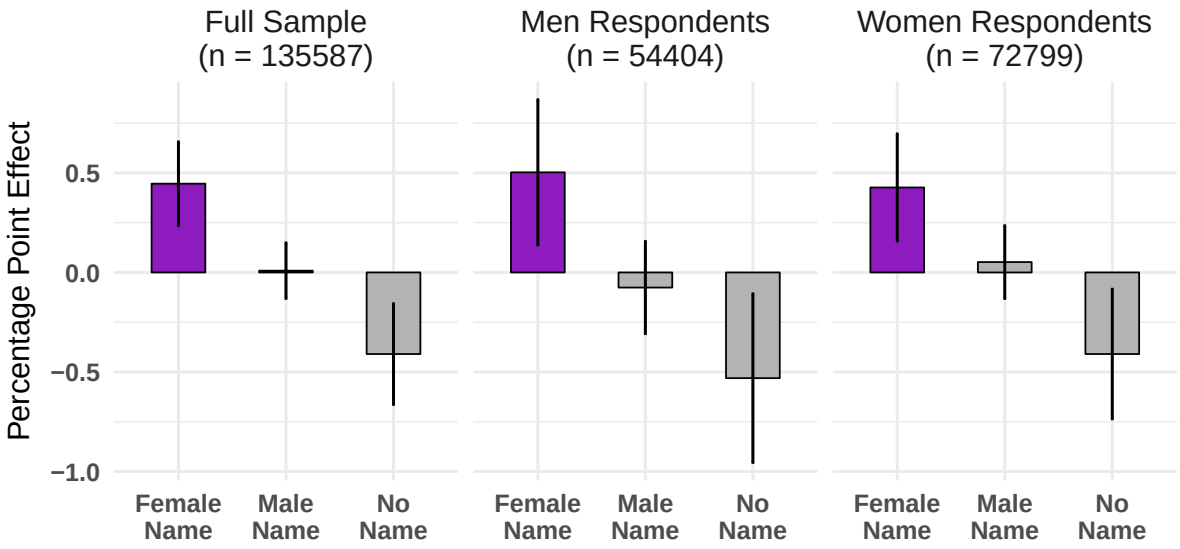

Figure 3: Mean Silencing. Figure shows the average treatment effect by treatment condition with 95% confidence intervals estimated using OLS. The comparison category is the ambiguous name condition. Data is from both studies. Table S13 displays the regression estimates depicted above.

Silencing Replies

We find that voters are more likely to silence (cut off) female-named volunteers than all other named conditions (Fig. 4). In the left-hand panel, we find that volunteers using female names were 0.446 percentage points more likely to be silenced than volunteers assigned an ambiguous name ($p < .001$). In contrast, volunteers assigned male names were about as likely to be silenced as the ambiguously-named volunteer. Unnamed volunteers were less likely to be silenced ($p < .01$). The middle and right-hand panels we again show little evidence of heterogeneity by respondent gender. The overall difference means that respondents force female-named volunteers to end 4.37 more conversations out of every 1000 text messages compared with male-named volunteers.

In a final analysis, we find no evidence that female-named volunteers were less effective

than volunteers assigned other name conditions. Although we would not expect a difference in efficacy because names were randomly assigned within volunteers, we assess whether poor treatment of volunteers correlates with lower efficacy—for instance, in generating commitments to call one's representative. In fact, we show the opposite is true: female-named volunteers obtain higher response rates and voter commitments to call their representatives than the other name conditions, despite their higher likelihood of receiving offensive, discouraging, and silencing responses (Tables S5-S7).

Discussion

Using experimental methods on a large sample of registered voters during real political campaigns, we find evidence that volunteers with female names will receive more offensive, discouraging and silencing replies than those with male and ambiguous names. These findings are consistent with research demonstrating that women are interrupted and harassed more. However, few prior studies have documented whether such findings held for direct experiences of political deliberation and conversation (*6*), and to our knowledge none examine real-life political activism.

The findings are striking given that our sampling frame is composed of individuals who previously shared their contact information with the organization. Our estimates therefore depict these behaviors within a reasonably "friendly" audience—that is, among ideologically like-minded voters who were willing to be contacted. We expect that the rates of uncivil behaviors will be much larger when the audience is not predisposed to be friendly. Indeed, many silencing responses come from respondents stating the organization has the wrong phone number, i.e., have not previously agreed to be contacted. However, rates of antisocial and sensitive behaviors may differ greatly depending on the medium of contact (e.g., face-to-face) used: the effects of perceived volunteer gender could be larger, smaller, or even reverse when campaigns

use other methods (*20*). Studying this variation seems crucial as many organizations report that their canvassers are predominantly women.

Understanding how voters react to different sorts of political volunteers and activists is important since deliberation underpins democracy (*21*). Voter-to-voter canvassing is one of the only methods proven to durably move voter attitudes on sensitive political issues (*22, 23*). Moreover, texting is among the few outreach tools available to campaigns—and increasingly common (*4*). Our findings suggest that it may matter a great deal who is texting (or calling, or door-knocking).

These studies offer a snapshot of political participants' treatment at one moment in time. As a given volunteer will send texts across all four name conditions, we cannot assess whether experiencing incivility influences participants' decisions to volunteer again. Moreover, those participating for the first time may find the possibility of experiencing incivility more intimidating than seasoned volunteers do. We hope that future research will explore these possibilities.

Finally, the findings add new dimension to a vast literature on civic engagement, political participation, and the mobilization of the public. While events like the 2017 Women's March and #MeToo movement demonstrate that more women are speaking up than ever before (*24*), we find that many individuals are not interested in listening.

References and Notes

1. D. C. Mutz, B. Reeves, *American Political Science Review* **99**, 1 (2005).

2. D. J. Brooks, J. G. Geer, *American Journal of Political Science* **51**, 1 (2007).

3. D. C. Mutz, *American Political Science Review* **101**, 621 (2007).

4. D. P. Green, A. S. Gerber, *Get out the vote: How to increase voter turnout* (Brookings Institution Press, 2019).

5. K. Manne, *Down girl: The logic of misogyny* (Oxford University Press, 2017).

6. C. F. Karpowitz, T. Mendelberg, *The silent sex: Gender, deliberation, and institutions* (Princeton University Press, 2014).

7. S. C. Herring, *The Information Society* **15**, 151 (1999).

8. S. Sobieraj, *Credible Threat: Attacks Against Women Online and the Future of Democratic Discourse* (Oxford University Press, Forthcoming).

9. M. Bertrand, S. Mullainathan, *American economic review* **94**, 991 (2004).

10. K. L. Schlozman, N. Burns, S. Verba, *The Journal of Politics* **61**, 29 (1999).

11. S. Verba, K. L. Schlozman, H. Brady, N. H. Nie, *British Journal of Political Science* **23**, 453 (1993).

12. J. S. Wong, S. K. Ramakrishnan, T. Lee, J. Junn, *Asian American political participation: Emerging constituents and their political identities* (Russell Sage Foundation, 2011).

13. E. M. Uslaner, M. Brown, *American politics research* **33**, 868 (2005).

14. K. L. Schlozman, N. Burns, S. Verba, J. Donahue, *American Journal of Political Science* pp. 267–293 (1995).

15. The ambiguously-gendered condition provides the best comparison group because it strictly varies the gender cue. The no-name condition acts as a different control condition by varying both the presence of a name and the gender cue associated with the name. Names may seem more personable and less like spam, generating friendlier responses. This is indeed what we find.

16. See Supplemental Materials for details on gender coding using the 'gender' package in R.

17. See Supplemental Materials for estimates that correct one-way non-compliance.

18. R. D. Putnam, *Bowling alone: The collapse and revival of American community* (Simon and Schuster, 2001).

19. T. Skocpol, M. P. Fiorina, *Civic engagement in American democracy* (Brookings Institution Press, 2004).

20. I. Krumpal, *Quality & Quantity* **47**, 2025 (2013).

21. A. De Tocqueville, *Democracy in america*, vol. 10 (Regnery Publishing, 2003).

22. D. Broockman, J. Kalla, *Science* **352**, 220 (2016).

23. J. L. Kalla, D. E. Broockman, *American Political Science Review* **112**, 148 (2018).

24. D. R. Fisher, D. M. Dow, R. Ray, *Science Advances* **3**, eaao1390 (2017).

25. See Supplemental Materials for examples of the SMS messages.

26. See Supplemental Materials for detailed coding instructions.

Supplementary Materials

Methods and Materials

Sampling Frame

In 2018, we worked with liberal political action committee NextGen America (NGA) to conduct two experiments evaluating whether female volunteers receive more harassing SMS responses that male volunteers. NGA was interested in this experiment because they had received comments from their female volunteers and staff that they felt harassed when sending get-out-the-vote (GOTV) SMS messages to voters. The organization wanted to find new ways to maintain the same programmatic efficacy while ensuring that their volunteers and staff members felt safe while participating in the political process.

The sample for the first experiment is 60,356 individuals who had interacted with the organization in the past; for the second experiment, 75,231 individuals drawn from the same population. The organization thought these individuals were more likely to attend a given political event because an internal model suggested that these individuals were likelier to vote for Democratic candidates.

Treatments

After creating the sample, individual respondents were randomly assigned to one of four conditions (*25*). Individuals in the first condition received a SMS message from a male-named volunteer named Michael. Individuals in the second condition received a message from a female-named volunteer named Jessica. Those in the third condition received a message from a volunteer with an ambiguously-gendered name, Taylor. Individuals in the last condition received messages with no name given (*15*). Volunteers did not know about the experiment's purpose. They were simply informed that the organization was trying something new to improve their

messaging program.

Dependent Variables

We measure our first dependent variable, *offensiveness*, in two ways. The measure in the first study is a self-report: how offensive volunteers felt the responses they received from each individual voter were. In the messaging system, after volunteers correspond with a voter, volunteers are asked whether the response they received was "offensive" or not. If the volunteer felt the message was offensive, then they marked "Yes." If the volunteer marked "Yes," then the respondent was coded as a 1, and if not, then they were coded as a 0. The measure of "offensiveness" in the second study is generated by ratings from two independent coders (*26*). Each coder rated the messages on a five-point scale from "non-offensive" (1), to "very offensive" (5). The inter-coder correlation was .82, which suggests that the measure is reasonably reliable. To compare scores between the two studies (main results reported in the paper), we also created a binary measure for Study 2: a message would be marked as offensive if either coder had marked it as offensive (1), and inoffensive if neither had (0). Our offensiveness findings for Study 2 hold whether the five-point scale is used or the binary scale is used, so for brevity we report only the binary analyses in the main paper, which allows us to condense the findings for both studies.

Our second dependent variable, *discouragement*, collected only in the second study, measures whether a message encourages or discourages a volunteer from volunteering again in the future. Two naive coders rated each message on a scale from "really encouraging" (1), to "really discouraging" (7). The correlation between the two coders was .76, which suggests moderate reliability. We again summed and normalized the two coders' ratings. To ensure our estimates of discouraging behavior are conservative, when a voter did not respond, we coded them as a 1 (really encouraging).

Our third dependent variable, *silencing*, measures whether a voter intimidates or instructs

the volunteer to cease further political discourse. Organization rules instruct volunteers to "opt-out" respondents from future contact if the respondent requests the volunteer to stop contacting them, or if the respondent harasses the volunteer. If the respondent was marked in the system as "opted-out," they were given as a 1, and if not, a 0. We treat this as a measure for "silencing" behavior because it prevents the volunteer from having any future conversation with the individual.

Additional Independent Variables

In addition to our treatment variables, we also estimate whether effects differ based on the respondent's gender, on the assumption that women and men might differ in their base propensity to send such messages as well as in their likelihood of responding one way to female volunteers and another to male volunteers.

We did not have gender information on respondents because some respondents did not originally provide their gender identity when providing the organization information. We used the genderizer package in R to code whether someone's name seemed male or female.

Inter-coder Reliability Statistics

The correlation between the two coders on the discouragement scale was .95; for the offensiveness scale, the correlation was .74.

We also tried replicating our results with another measure of toxicity, via PerspectiveAPI. PerspectiveAPI provides toxicity scores for comments on social media and news articles, among other things, using an algorithm to automatically classify texts based on their similarity to known toxic language. Our hope in using it was that we would be able to compare "apples to apples" across the two studies—a beneficial addition because they currently rely on different human coders. However, we found that the scores did not correlate well with any of our

measures, perhaps because our snippets of text (SMS messages) are so short and thus subject to more measurement error.

For experiment 1, PerspectiveAPI toxicity scores correlated with our coder-rated offensiveness measure at .56 and silencing measure at .45. For experiment 2, toxicity scores correlated with our coder-rated offensiveness measure at .62, discouragement measure at .61, and silencing measure at .47.

Ultimately, we were not confident that the toxicity scores would be informative because of their low correlations with our existing measures. Indeed, we relied on codings by human volunteers from the beginning due to pre-existing concerns about the difficulty of using automated classifiers (which typically measure similarity between two strings, and thus are more accurate the longer the strings are). Accordingly, we found ourselves disappointed but unsurprised that we could not use PerspectiveAPI toxicity scores as a way to externally validate our volunteers' ratings.

Next, we checked whether offensiveness, silencing, and discouragement were measuring different concepts. In experiment 1, the correlation between volunteer-coded offensiveness and silencing was 0.20. We felt confident that we had measured unique concepts. In experiment 2, the correlation between coder-rated offensiveness and silencing remained low at 0.33. The correlation between coder-rated offensiveness and coder-rated discouragement was 0.5. We expected these concepts to have a slightly higher correlation, so this is unsurprising, but still suggest differences between the concepts. However, the correlation between discouragement and silencing was higher than we would have liked, at 0.76. This suggests a strong overlap between the two concepts.

Study Context, Delivery, and Timing

NextGen America is a national political action committee that focuses on engaging young people in the political process to support progressive causes. They typically organize on and around college campuses focusing on voter registration and voter turnout. They were originally founded in 2013 as NextGen Climate and primarily focused on environmental issues.

The first experiment was conducted during NGA's response to the Marjory Stoneman Douglas High School shooting in Florida. Activists were organizing an event, March for our Lives, to protest the lack of gun control legislation. NGA texted their list of progressive voters to encourage them to participate in the protest. The organization has recruited for other protests in the past by supporting the Women's March and the Climate March.

In the first experiment, the organization failed to deliver 18,457 SMS messages out of those individuals assigned to be part of the study, because they lacked sufficient volunteers. The organization failed to deliver 36,290 SMS messages in the second experiment for the same reason.

In the second experiment, NGA texted progressive voters to encourage them to lobby their members of Congress to call for then-Environmental Protection Agency Director Scott Pruitt's removal. Pruitt had recently been accused of several ethical violations, which sparked intense lobbying for his removal by progressive organizations. NGA has recruited voters to lobby members of Congress in the past on several political issues.

Three days before the March for our Lives protest in March, volunteers began sending messages to the first experiment's sample through Relay, an online SMS messaging platform. Later in April, volunteers began texting the second experiment's sample through Relay. Volunteers were not told about the study's intent to reduce potential demand effects. Additionally, only a select few staffers in the organization knew about the study in order to reduce the number of people who could induce demand effects.

Coding Instructions

You will be given individual response text messages from people who originally received text messages from NGA. You will undertake two coding tasks for each text:

Task 1: How offensive does each text seem? Each message is to be coded on a 1 to 5 scale where 1 is non-offensive and 5 is very offensive. Below illustrates what the scale looks like:

1. Non-offensive
2. Slightly offensive
3. Moderately offensive
4. Fairly offensive
5. Very offensive

Task 2: Would this message encourage or discourage you from volunteering in the future? Each message is to be coded on a 1 to 7 scale where 1 is really encourage and 7 is really discourage. Below illustrates what the scale looks like:

1. Really encourage
2. Somewhat encourage
3. Slightly encourage
4. Neither encourage nor discourage
5. Slightly discourage
6. Somewhat discourage

7. Really discourage

Each of you will be given the same list of text messages. However, you should not consult each other while coding because this could bias the results. Your codings don't have to match each other's because we will make an average "rating" from your codings, so it's not an error if you disagree.

If either of you have a question or comment, let me know and I'll help you on the individual text message. Some text messages may be broken up because of processing errors on Relay's side; if you find this, let me know and I'll help you.

Sample Messages

First Experiment Message

Hi [VOTER'S NAME], I'm [EXPERIMENTAL TREATMENT NAME], a volunteer with NextGen America. This Saturday, March 24, young people have organized a march to spread awareness and demand gun safety legislation from our elected officials. Find a local March For Our Lives here: http://nxtgn.us/dkh - Can we count on you to join us at the march?

Second Experiment Message

Hey [VOTER'S NAME], it's [EXPERIMENTAL TREATMENT NAME] volunteering w/ NextGen America. EPA Administrator Scott Pruitt is under investigation for his waste of taxpayer dollars and shady ties to corporate polluters. Urge your members of Congress to call on Scott Pruitt to resign: http://NXTGN.US/du5

Response Rates

Study 1 Response Rates

Table S1: Response Rate by Experimental Condition, Study 1

Experimental Condition	Response Rates
Female Name	27.57%
Ambiguous Name	26.46%
Male Name	26.14%
No Name	24.78%

Table S2: Response Rate by Experimental Condition and Respondent Gender, Study 1

Experimental Condition	Respondent Gender	Response Rates
Female Name	Men	28.94%
Female Name	Women	27.59%
Ambiguous Name	Men	29.16%
Ambiguous Name	Women	25.21%
Male Name	Men	27.07%
Male Name	Women	26.35%
No Name	Men	25.23%
No Name	Women	24.96%

Study 2 Response Rates

Table S3: Response Rate by Experimental Condition, Study 2

Experimental Condition	Response Rates
Female Name	8.44%
Ambiguous Name	8.08%
Male Name	8.47%
No Name	7.33%

Table S4: Response Rate by Experimental Condition and Respondent Gender, Study 2

Experimental Condition	Respondent Gender	Response Rates
Female Name	Men	9.07%
Female Name	Women	8.14%
Ambiguous Name	Men	8.83%
Ambiguous Name	Women	7.35%
Male Name	Men	8.5%
Male Name	Women	7.84%
No Name	Men	8.09%
No Name	Women	6.86%

In the full sample, female-named volunteers received more responses, while male-named volunteers produced little change. Unnamed volunteers were substantially less likely to receive a reply. In the women-only sample, female-named male-named volunteers received more replies, but only the former coefficient was statistically significant. The unnamed volunteer was again less likely to receive replies, but the effect was not statistically significant. Finally, in the men-only sample, female-named volunteer were again received more replies, but this effect was not statistically significant. Male-named and unnamed volunteers received fewer replies, but only the latter was statistically significant.

Table S5: Effect of a Volunteer's Perceived Gender on the Likelihood of Receiving a Response, Both Studies

	Responses		
	Full Sample	Women only	Men only
Female Name	0.422^{*}	0.568^{*}	0.344
	(0.178)	(0.237)	(0.296)
Male Name	0.017	0.300	-0.386
	(0.131)	(0.175)	(0.217)
No Name	-0.907^{***}	-0.365	-1.751^{***}
	(0.239)	(0.325)	(0.386)
Constant	4.707^{***}	9.247	4.520^{**}
	(0.888)	(0.000)	(1.562)
N	135,587	72,799	54,404
R^2	0.043	0.010	0.047

$^{\dagger}p < .1$; $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$

Volunteer-level and experiment-level fixed effects are included. Robust standard errors are reported in the parentheses.

Texts Sent

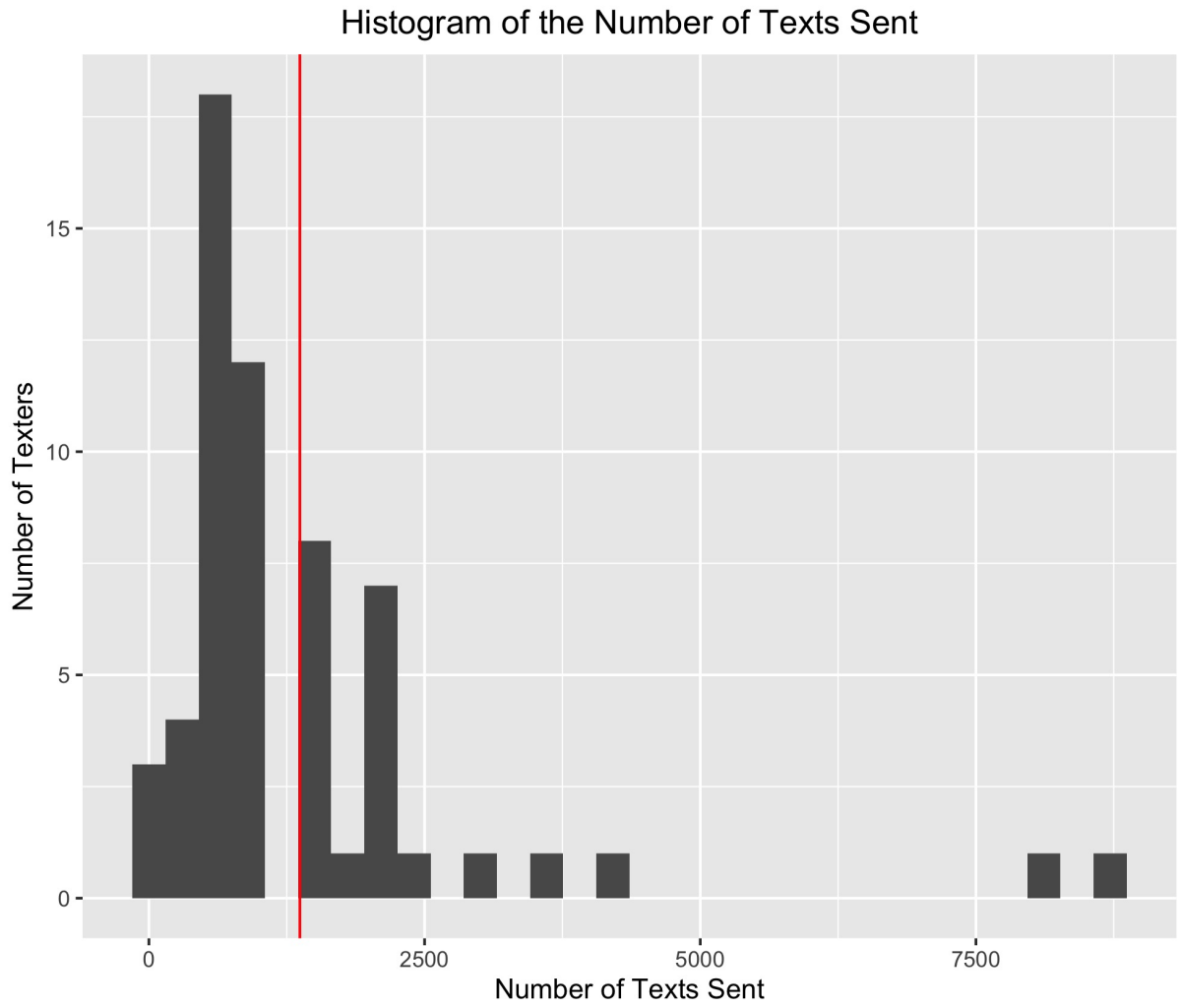

Figure S1: Number of texts sent. Figure shows how many texts were sent by volunteers across both studies. The red line indicates that on average volunteers sent 1,369 texts.

Regression Analyses for Stated March Attendance and Constituent Calling

Study 1

In the full sample, unnamed volunteers received fewer commitments to attend the Women's March, however, this effect was statistically insignificant. The female-named and male-named conditions' coefficients were substantively small. In the women-only sample, the female-named condition produced little change in reported march attendance. The male-named volunteers received more attendance commitments, while the unnamed volunteers statistically significantly fewer commitments. Finally, female-named and unnamed volunteers were less likely to obtain march attendance commitments, while the male-named volunteers produced little change.

Table S6: Effect of a Volunteer's Perceived Gender on the Likelihood of March Attendance, Using Self-Reports

	Reported March Attendance		
	Full sample	Women only	Men only
Female Name	−0.034	0.082	−0.210
	(0.189)	(0.073)	(0.251)
Male Name	0.079	0.224	−0.010
	(0.132)	(0.147)	(0.268)
No Name	−0.328	$-0.354^{\dagger}$	−0.321
	(0.162)	(0.210)	(0.240)
Constant	0.321^{*}	0.151	0.613^{*}
	(0.132)	(0.147)	(0.268)
N	60,356	31,656	25,083
R^2	0.006	0.008	0.005

$^{\dagger}p < .1$; $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$

Data derived from Study 1. Baseline category is the ambiguously-gendered name condition. Volunteer-level fixed effects are included but not reported. Robust-clustered standard errors are reported in the parentheses.

Study 2

Across all three samples, female-named volunteers obtained more commitments from respondents to call their congresspersons. Male-named and unnamed volunteers produced little change in commitments across all three samples.

Table S7: Effect of a Volunteer's Perceived Gender on the Likelihood of Calling Their Representative, Using Self-Reports

	Reported Calling		
	Full sample	Women only	Men only
Female Name	0.076^{***}	0.029^{***}	0.143^{***}
	(0.163)	(0.004)	(0.001)
Male Name	-0.005	-0.001	0.0004
	(0.006)	(0.001)	(0.0004)
No Name	-0.005	-0.001	0.0004
	(0.006)	(0.001)	(0.0004)
Constant	0.605^{***}	0.364^{***}	0.847^{***}
	(0.006)	(0.001)	(0.0004)
N	75,231	41,143	29,321
R^2	0.003	0.004	0.004

$^{\dagger}p < .1$; $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$

Data derived from Study 2. Baseline category is the ambiguously-gendered name condition. Volunteer-level fixed effects are included but not reported. Robust-clustered standard errors are reported in the parentheses.

Regression Analyses for Results Reported Visually

The following tables provide the raw regression results for the analyses reported in Figures 1-3 in the paper. We first provide the results of the offensiveness meta-analysis (both studies) reported in the paper, and then provide the results for each study individually.

Offensiveness Across Both Studies

Table S8: Effect of a Volunteer's Perceived Gender on the Likelihood of Receiving an Offensive Message, Both Studies

	Offensive Replies		
	Full Sample	Women only	Men only
Female Name	0.177^{***}	0.202^{***}	0.172^{*}
	(0.042)	(0.051)	(0.075)
Male Name	-0.097^{**}	-0.053	-0.144^{*}
	(0.031)	(0.036)	(.057)
No Name	0.148^{***}	0.139^{**}	0.181^{*}
	(.041)	(.048)	(.075)
Constant	0.232^{***}	0.156^{***}	0.317^{***}
	(.999)	(1.373)	(1.672)
N	135,587	72,799	54,404
Adjusted R^2	0.011	0.010	0.013

$^{\dagger}p < .1$; $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$

Baseline category is the ambiguously-gendered name condition. Experiment-level fixed effects are included but not reported. Robust standard errors are reported in the parentheses.

In both experiments, we find that volunteers assigned female names receive more offensive responses from voters. This table presents regression results overall and by gender of the respondent, with a dummy variable for whether the data comes from the first experiment. The first column reports the regression results by treatment condition for the full sample. The second column reports results for women respondents; the third, for men respondents.

We see across all three columns that female-named volunteers are significantly more likely to receive offensive messages than ambiguously-named volunteers, as are unnamed volunteers, while male-named volunteers receive significantly fewer offensive messages than ambiguously-named volunteers except among female respondents (where the estimate is still negative, but the $p = .137$). In all three samples, female-named volunteers receive offensive messages significantly more than male-named volunteers.

Study 1

In the first experiment, we find evidence that volunteers assigned female names self-report more offensive responses from voters. This table presents regression results overall and by gender of the respondent. The first column reports the regression results by treatment condition for the full sample. The second column reports results for women respondents; the third, for men respondents.

Table S9: Effect of a Volunteer's Perceived Gender on the Likelihood of Receiving an Offensive Message, Using Self-Reports

	Offensive Replies		
	Full sample	Women only	Men only
Female Name	0.151^{*}	0.176^{*}	0.158
	(0.060)	(0.073)	(0.132)
Male Name	-0.013	0.029	-0.054
	(0.049)	(0.060)	(0.092)
No Name	0.074	0.078	0.095
	(0.055)	(0.064)	(0.143)
Constant	0.191^{***}	$.125^{**}$	0.271^{***}
	(0.035)	(0.040)	(0.092)
N	60,356	31,656	25,083
Adjusted R^2	0.001	0.001	0.001

$^{\dagger}p < .1$; $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$

Data derived from Study 1. Baseline category is the ambiguously-gendered name condition. Robust standard errors are reported in the parentheses.

Overall, we find that respondents in Study 1 were more likely to send offensive messages to female-named volunteers than to ambiguously-named volunteers (two-tailed $p = .011$). In the second-column, we see that this is also true when female-named volunteers contact female respondents. While in the third column, for male respondents, the estimate drops from statis-

tical significance, the estimate is not significantly different from those for women respondents, and we again find that the difference between the treatment of female-named and male-named volunteers is significant among male respondents ($p = .038$).

We also conduct an analysis of Study 1 that employs fixed effects for individual volunteers. This allows us to eliminate the possibility that our effects are driven by differences in individual volunteers' propensity to rate messages as offensive. In this analysis, we find that volunteers assigned to use female names are 0.113 percentage points more likely than ambiguously-named volunteers to receive offensive messages during a campaign (two-tailed $p = .097$). In contrast, male-named volunteers are no more likely to experience additional offensive messages, and unnamed volunteers were slightly (but not significantly) more likely to receive offensive replies. To contextualize the size of these effects, this means that for every 1,000 messages a female-named volunteer sends, she receives 1.13 more offensive messages than a texter with an ambiguously-gendered name.

To examine potential heterogeneity among voters, in columns two and three we subset to voters who we could identify as women and men. When texting women, the penalty for being female-named relative to an ambiguous name decreases to 0.07 percentage points; when texting men, it increases to 0.183, though in neither case is it statistically significant. Ostensibly male volunteers saw the opposite pattern (more offensive replies from women and fewer from men), but again the pattern is not statistically significant.

Table S10: Effect of a Volunteer's Perceived Gender
on the Likelihood of Receiving an Offensive Message,
Using Self-Reports,
with Volunteer Fixed Effects

	Offensive Replies		
	Full sample	Women only	Men only
Female Name	$0.113^{\dagger}$	0.071	0.183
	(0.068)	(0.073)	(0.132)
Male Name	-0.031	0.040	-0.112
	(0.051)	(0.063)	(0.092)
No Name	0.058	0.053	0.070
	(0.075)	(0.080)	(0.143)
Constant	0.231	-0.040	0.112
	(0.207)	(0.063)	(0.092)
N	60,356	31,656	25,083
Adjusted R^2	0.001	0.001	0.001

$^{\dagger}p < .1$; $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$

Data derived from Study 1. Baseline category is the ambiguously-gendered name condition. Volunteer-level fixed effects are included but not reported. Robust standard errors are reported in the parentheses.

Study 2

The next table shows the regression results for Study 2, which used naive raters' codings of offensiveness. In the first column, we show that female-named and unnamed volunteers were significantly more likely to receive offensive replies than an ambiguously-named volunteer. Male-named volunteers were slightly less likely to receive offensive replies than the ambiguously-named volunteer. Looking at columns two and three, we see that female-named volunteers fare slightly worse among women voters than they do among men voters, but the estimate remains substantively unchanged. Male-named volunteers fare equally well with men and women voters. Unnamed volunteers are treated quite differently by men and women: they are much more likely to receive offensive content from men than from women.

Table S11: Effect of a Volunteer's Perceived Gender on the Likelihood of Receiving an Offensive Message, as Coded by Other Volunteers

	Offensiveness		
	Full Sample	Women only	Men only
	(1)	(2)	(3)
Female Name	0.040^{*}	0.052^{*}	0.034
	(0.020)	(0.024)	(0.037)
Male Name	-0.022	-0.023	-0.027
	(0.018)	(0.017)	(0.036)
No Name	0.094^{***}	0.050^{*}	0.148^{**}
	(0.026)	(0.024)	(0.053)
Constant	0.062^{***}	0.040^{**}	0.093^{***}
	(0.012)	(0.013)	(0.024)
N	75,231	41,143	29,321
Adjusted R^2	0.0003	0.0003	0.0005

$^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$

Robust standard errors are reported in parentheses.

The next table presents the same analysis as above using the dichotomized volunteer-coded

offensiveness measure. The first column show that female-named and unnamed volunteers were significantly more likely to receive offensive replies than ambiguously-named volunteers. Male-named volunteers were significantly less likely to receive offensive messages than ambiguously-named volunteers. We repeat the same analysis with columns two and three. Like before, female-named and unnamed volunteers were significantly more likely to receive offensive replies, while male-named volunteers were significantly less likely to receive offensive replies.

Table S12: Effect of a Volunteer's Perceived Gender on the Likelihood of Receiving an Offensive Message, as Coded by Other Volunteers and Dichotomized

	Offensiveness		
	Full Sample	Women only	Men only
	(1)	(2)	(3)
Female Name	0.199^{**}	0.223^{**}	$0.186^{\dagger}$
	(0.059)	(0.072)	(0.106)
Male Name	-0.164^{***}	-0.115^{**}	-0.222^{**}
	(0.039)	(0.043)	(0.075)
No Name	0.208^{***}	0.186^{**}	0.254^{*}
	(0.059)	(0.069)	(0.109)
Constant	0.228^{***}	0.155^{***}	0.315^{***}
	(0.035)	(0.039)	(0.066)
N	75,231	41,143	29,321
R^2	0.0008	0.0008	0.0009

$^{\dagger}p < .1$; $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$

Robust standard errors are reported in parentheses.

We also find evidence that female-named volunteers receive more discouraging replies. In the next table, we regress ratings of how discouraging messages were on the experimental conditions. In the first column, we show that female-named and unnamed volunteers were significantly more likely to receive discouraging replies than an ambiguously-named volunteer, while

male-named volunteers were significantly less likely to receive discouraging replies. Female-named volunteers are slightly but not significantly more likely to receive a discouraging reply from a woman than a man (columns two and three), while the opposite is true for male-named volunteers. Unnamed volunteers received significantly more discouraging replies from men than from women.

Table S13: Effect of a Volunteer's Perceived Gender on the Likelihood of Receiving a Discouraging Message, as Coded by Other Volunteers

	Discouraging		
	Full Sample	Women only	Men only
	(1)	(2)	(3)
Female Name	0.393^{***}	0.438^{***}	0.374^{***}
	(0.067)	(0.090)	(0.110)
Male Name	-0.355^{***}	-0.336^{***}	-0.379^{***}
	(0.048)	(0.062)	(0.080)
No Name	0.413^{***}	0.350^{***}	0.521^{***}
	(0.068)	(0.086)	(0.116)
Constant	0.554^{***}	0.519^{***}	0.591^{***}
	(0.041)	(0.053)	(0.067)
N	75,231	41,143	29,321
Adjusted R^2	0.003	0.003	0.003

$^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$

Robust standard errors are reported in parentheses.

In both studies, we find strong evidence that voters are more likely to silence female-named volunteers than all other named conditions. In the final table in this section, we report the OLS regression results with an added dummy variable indicating whether the data comes from the first or second experiment. In the first column, we find that volunteers using female names were 0.446 percentage points more likely to receive silencing replies from voters than volunteers assigned an ambiguous name. In contrast, volunteers assigned male names were about as likely

Table S14: Effect of a Volunteer's Perceived Gender on the Likelihood of Receiving a Silencing Message, Both Studies

	Silencing Replies		
	Full Sample	Women only	Men only
Female Name	0.446^{***}	0.427^{**}	0.503^{**}
	(0.107)	(0.137)	(0.186)
Male Name	0.009	0.052	-0.076
	(0.071)	(0.093)	(.118)
No Name	-0.410^{**}	-0.410^{*}	-0.531^{*}
	(.129)	(.166)	(.216)
Constant	4.904^{***}	4.951^{***}	5.369^{**}
	(.999)	(1.373)	(1.672)
N	135,587	72,799	54,404
Adjusted R^2	0.011	0.010	0.013

$^{\dagger}p < .1$; $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$

Volunteer-level and experiment-level fixed effects are included. Robust standard errors are reported in the parentheses.

to be silenced as the ambiguously-named volunteer. Unnamed volunteers were less likely to be silenced. The difference means that female-named volunteers are forced to end 4.45 more conversations out of every 1000 text messages they send to voters than male-named volunteers do.

The results in columns two and three of this table show little evidence of heterogeneity in silencing behavior amongst men and women. Ostensibly female volunteers are slightly but not significantly more likely to receive a silencing reply from men than from women. For male-named volunteers, women are more likely to silence than men, but the estimates are again not statistically distinguishable from one another. Unnamed volunteers were less likely to be silenced by women than by men. In sum, even after accounting for key differences in men and women respondents' propensity to cut off conversations, we find that volunteers assigned female names were more likely to have their conversations discontinued, even after accounting

for differences in the propensity of men and women voters to cut off conversations.

Instrumental Variables Analyses

In the following section, we provide estimates of the average treatment effect on the treated, using instrumental variable regressions to correct for one-sided noncompliance.

IV: Self-Reports of Offensiveness, Study 1

The first column shows that when a voter receives a message from a volunteer with a female name, male name, and no name, they are .19 percentage points more likely, .05 percentage points less likely, and .09 percentage points more likely to respond with an offensive reply. None of these results are statistically significant below the 5% threshold. However, the point estimates are fairly stable when we subset to women only (column two) and men only (column three).

Table S15: Complier Effect of a Volunteer's Perceived Gender on the Likelihood of Receiving an Offensive Message, Study 1

	Offensive Replies		
	Full sample	Women only	Men only
	(1)	(2)	(3)
Female Name	0.191	0.125	0.307
	(0.119)	(0.127)	(0.232)
Male Name	−0.054	0.069	−0.193
	(0.090)	(0.111)	(0.161)
No Name	0.087	0.089	0.095
	(0.094)	(0.104)	(0.179)
Constant	0.254	−0.069	0.193
	(0.219)	(0.111)	(0.161)
N	60,356	31,656	25,083
Adjusted R^2	0.001	0.001	0.001

$^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$

Volunteer-level fixed effects are included. Robust-clustered standard errors are reported in the parentheses.

IV: Offensiveness as Rated by Naive Raters, Study 2

The first column shows that when a voter receives a message from a volunteer with a female name or no name, they were .06 percentage points more likely and .12 percentage points more likely to respond with an offensive reply, respectively. When subset to women (column two) and men (column three) respondents, our results replicate those in the paper and become stronger, except for the female name coefficient in the men only sample.

Table S16: Complier Effect of a Volunteer's Perceived Gender on the Likelihood of Receiving an Offensive Message, as Coded by Other Volunteers

	Offensiveness		
	Full Sample	Women only	Men only
	(1)	(2)	(3)
Female Name	0.055*	0.071*	0.047
	(0.027)	(0.033)	(0.051)
Male Name	−0.165	−0.169	−0.207
	(0.135)	(0.124)	(0.277)
No Name	0.121***	0.064*	0.189**
	(0.033)	(0.031)	(0.068)
Constant	0.062***	0.040**	0.093***
	(0.012)	(0.013)	(0.024)
N	75,231	41,143	29,321
Adjusted R^2	0.00003	−0.0001	0.0001

*p < .05; **p < .01; ***p < .001

Robust standard errors are reported in parentheses.

IV: Dichotomized Offensiveness, Both Studies

This table shows the IV analysis for how respondents replies to differently gendered names using the dichotomized offensive ratings from both studies. The first column shows that female-named and unnamed volunteers were significantly likelier than ambiguously-named volunteers to receive offensive replies. Male-named volunteers were significantly less likely to receive offensive replies than ambiguously-named volunteers. The second and third columns repeat the analysis from before with women-only and men-only samples, respectively. They both show the same pattern where female-named and unnamed volunteers are significantly more likely to receive offensive replies, while male-named volunteers are significantly less likely to receive offensive replies.

Table S17: Complier Effect of a Volunteer's Perceived Gender on the Likelihood of Receiving an Offensive Message, Both Studies

	Offensive Replies		
	Full Sample	Women only	Men only
Female Name	0.240^{***}	0.272^{***}	0.235^{*}
	(0.057)	(0.069)	(0.102)
Male Name	-0.287^{**}	-0.160	-0.412^{*}
	(0.091)	(0.108)	(0.164)
No Name	0.191^{***}	0.179^{**}	0.233^{*}
	(0.053)	(0.062)	(0.097)
Constant	0.217^{***}	0.149^{***}	0.294^{***}
	(0.025)	(0.028)	(0.046)
N	135,587	72,799	54,404
R^2	0.0005	0.010	0.0003

$^{\dagger}p < .1$; $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$

Baseline category is the ambiguously-gendered name condition. Experiment-level fixed effects are included but not reported. Robust standard errors are reported in the parentheses.

IV: Discouragement as Rated by Naive Raters, Study 2

The first column shows that when a voter receives a message from a volunteer with a female name, male name, or no name, they were .54 percentage points more likely, 2.67 percentage points less likely, and .53 percentage points less likely to respond with a discouraging reply, respectively. When subset to women (column two) and men (column three) respondents, our results replicate the findings in the main paper.

Table S18: Complier Effect of a Volunteer's Perceived Gender on the Likelihood of Receiving an Discouraging Message, as Coded by Other Volunteers

	Discouraging		
	Full Sample	Women only	Men only
	(1)	(2)	(3)
Female Name	0.537***	0.598***	0.511***
	(0.091)	(0.122)	(0.150)
Male Name	−2.668***	−2.503***	−2.870***
	(0.368)	(0.472)	(0.620)
No Name	0.529***	0.446***	0.667***
	(0.086)	(0.109)	(0.148)
Constant	0.554***	0.519***	0.591***
	(0.041)	(0.053)	(0.067)
N	75,231	41,143	29,321
Adjusted R^2	−0.006	−0.005	−0.007

*p < .05; **p < .01; ***p < .001

Robust standard errors are reported in parentheses.

IV, Silencing, Meta-Analysis

The first column shows that when a voter receives a message from a volunteer with a female name or no name, they were .95 percentage points more likely and .55 percentage points less likely to respond with a silencing reply, respectively. When subset to women (column two) and men (column three) respondents, our results replicate the previous results and become stronger.

Table S19: Complier Effect of a Volunteer's Perceived Gender on the Likelihood of Receiving a Silencing Message, Both Studies

	Silencing Replies		
	Full Sample	Women only	Men only
Female Name	0.947^{***}	0.923^{**}	1.024^{*}
	(0.249)	(0.320)	(0.427)
Male Name	0.024	0.183	-0.267
	(0.251)	(0.336)	(0.402)
No Name	-0.553^{*}	-0.529^{*}	-0.795^{*}
	(0.256)	(0.333)	(.418)
Constant	-0.078^{***}	-0.060^{**}	-0.116^{***}
	(0.019)	(0.026)	(0.028)
N	135,587	72,799	54,404
R^2	0.012	0.011	0.014

$^{\dagger}p < .1$; $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$

Volunteer-level and experiment-level fixed effects are included. Robust standard errors are reported in the parentheses.

Robustness Check of Dichotomized Offensive Coding

The analyses below codes a message as offensive when the offensive index measure is greater than 50, 40, 30, and 20. We show that our results are broadly consistent with our main results and are not sensitive to our coding decisions. Across all four columns, female-named and unnamed volunteers were significantly more likely to receive an offensive messages compared to the ambiguously-named volunteer. Across all four columns, male-named volunteers were less likely to receive an offensive message compared to ambiguously-named volunteers, but none of these effects were statistically significant.

Table S20: Effect of a Volunteer's Perceived Gender on the Likelihood of Receiving an Offensive Message, Both Studies - Robustness Check

	Offensive Replies			
	> 50	> 40	> 30	> 20
Female Name	0.076^{**}	0.076^{**}	0.076^{*}	0.076^{*}
	(0.029)	(0.029)	(0.031)	(0.033)
Male Name	-0.009	-0.009	-0.021	-0.038
	(0.025)	(0.025)	(0.026)	(0.027)
No Name	0.086^{**}	0.086^{**}	0.083^{**}	0.086^{**}
	(0.030)	(0.030)	(0.031)	(0.033)
Constant	$0.031^{\dagger}$	$0.031^{\dagger}$	0.055^{**}	0.094^{***}
	(0.018)	(0.018)	(0.020)	(0.022)
N	135,587	135,587	135,587	135,587
R^2	0.0006	0.0006	0.0005	0.0003

$^{\dagger}p < .1$; $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$

Baseline category is the ambiguously-gendered name condition. Experiment-level fixed effects are included but not reported. Robust standard errors are reported in the parentheses.

Regression Analyses Conditioning on Response

The following tables describe how respondents responded to differently named volunteers conditional on responding in the first place. We refrain from providing these results in the main text because they are vulnerable to post-treatment bias: we know that people are already more likely to respond to female-named volunteers than male-named volunteers. Therefore, the only way we can interpret these results is if the sequential ignorability assumption holds, which is both unlikely and unverifiable. However, we provide these analyses to provide evidence that our results are not purely a function of the fact that female-named volunteers are more likely to receive responses. These analyses require us to drop a substantial portion of our sample, which will inflate our standard errors.

Table S21: Effect of a Volunteer's Perceived Gender on the Likelihood of Receiving a Silencing Message Conditional on Receiving a Message, Both Studies

	Silencing Replies		
	Full Sample	Women only	Men only
Female Name	2.132	3.547	1.245
	(1.838)	(2.535)	(2.798)
Male Name	−2.542	−2.192	−1.741
	(2.037)	(2.824)	(3.084)
No Name	−2.955	−3.906	−2.731
	(1.850)	(2.547)	(2.821)
Constant	49.808***	53.515***	46.185***
	(1.992)	(2.762)	(3.081)
N	4,946	2,561	2,169
R^2	0.030	0.052	0.013

$^{\dagger}p < .1$; $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$

The ambiguous-name condition is the omitted (comparison) category. Experiment-level fixed effects are included. Robust standard errors are reported in the parentheses.

The first column presents the effect of a gendered name on the likelihood of silencing con-

ditional on receiving a reply for the full sample. Female-named volunteers are 2.132 percentage points more likely to receive silencing replies. Male-named and unnamed volunteers were 2.542 percentage points and 2.955 percentage points less likely to receive silencing replies, respectively. For context, if a female-named volunteer were to receive 100 replies, they would receive 2.13 more silencing replies. None of these effects are statistically significant.

The second column presents the effect of a gendered name on the likelihood of silencing conditional on receiving a reply for only women respondents. Female-named volunteers were 3.547 percentage points more likely to receive silencing replies. Male-named and unnamed volunteers were 2.192 percentage points and 3.906 percentage points less likely to receive a silencing reply, respectively. For context, if a female-named volunteer were to receive 100 replies, they would receive 3.55 more silencing replies. None of these effects are statistically significant.

The third column displays the effect of a gendered name on the likelihood of silencing conditional on receiving a reply for only men respondents. Female-named volunteers were 1.245 percentage points more likely to receive a silencing reply. Male-named and unnamed volunteers were 1.741 percentage points and 2.731 percentage points less likely to receive silencing replies. None of these effects were statistically significant.

The first column displays the effect of a gendered name on the likelihood of receiving an offensive reply conditional on receiving a reply for the full sample. Female-named volunteers were 1.083 percentage points more likely to receive offensive replies. Male-named volunteers were 0.867 percentage points more likely to receive offensive replies. Unnamed volunteers were 1.347 percentage points more likely to receive offensive replies. Only the female-named volunteer condition was statistically significant $(p < .1)$.

The second column presents the effect of a gendered name on the likelihood of receiving an offensive reply conditional on receiving a reply for women respondents. Female-named volun-

Table S22: Effect of a Volunteer's Perceived Gender on the Likelihood of Receiving a Offensive Message Conditional on Receiving a Message, Both Studies

	Offensive Replies		
	Full Sample	Women only	Men only
Female Name	$1.083^{\dagger}$	$1.679^{\dagger}$	.871
	(0.628)	(0.865)	(0.887)
Male Name	0.867	0.710	0.709
	(1.016)	(1.081)	(1.637)
No Name	1.347	0.965	2.010
	(0.828)	(0.836)	(1.208)
Constant	5.666^{***}	3.578^{***}	7.856^{***}
	(0.704)	(0.725)	(1.119)
N	4,946	2,561	2,169
R^2	0.004	0.003	0.008

$^{\dagger}p < .1$; $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$

The ambiguous-name condition is the omitted (comparison) category. Experiment-level fixed effects are included. Clustered-robust standard errors are reported in the parentheses.

teers were 1.679 percentage points more likely to receive offensive replies. Male-named volunteers were .71 percentage points more likely to receive offensive replies. Unnamed volunteers were .965 percentage points more likely to receive offensive replies. Only the female-named volunteer condition was statistically significant $(p < .1)$.

The third column shows the effect of a gendered name on the likelihood of receiving an offensive reply conditional on receiving a reply for men respondents. Female-named volunteers were .871 percentage points more likely to receive offensive replies. Male-named volunteers were .709 percentage points more likely to receive offensive replies. Unnamed volunteers were 2.01 percentage points more likely to receive offensive replies. None of these effects were statistically significant.

Table S23: Effect of a Volunteer's Perceived Gender on the Likelihood of Receiving a Discouraging Message Conditional on Receiving a Message, Second Study

	Discouraging Replies		
	Full Sample	Women only	Men only
Female Name	1.140	2.466	−1.561
	(1.760)	(2.344)	(2.767)
Male Name	8.022**	3.935	10.234*
	(2.805)	(3.674)	(4.659)
No Name	0.901	−1.630	1.339
	(1.787)	(2.401)	(2.578)
Constant	47.247***	47.917***	48.408***
	(1.401)	(1.914)	(2.120)
N	4,946	2,561	2,169
R^2	0.008	0.010	0.016

$^{\dagger}p < .1$; $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$

The ambiguous-name condition is the omitted (comparison) category. Robust standard errors are reported in the parentheses.

The first column presents the effect of gendered names on the likelihood of a discouraging

reply conditional on receiving a reply. Female-named volunteers were 1.14 percentage points more likely to receive a discouraging reply. Male-named volunteers were 8.022 percentage points more likely to receive a discouraging reply. Unnamed volunteers were .901 percentage points more likely to receive a discouraging reply. Only the male-named effect was statistically significant $(p < .01)$.

The second column presents the effect of gendered names on the likelihood of a discouraging reply conditional on receiving a reply for women respondents. Female-named volunteers were 2.466 percentage points more likely to receive discouraging replies. Male-named volunteers were 3.935 percentage points more likely to receive discouraging replies. Unnamed volunteers were 1.63 percentage points less likely to receive discouraging replies. None of these effects were statistically significant.

The third column presents the effect of gendered names on the likelihood of a discouraging reply conditional on receiving a reply for men respondents. Female-named volunteers were 1.561 percentage points less likely to receive a discouraging reply. Male-named volunteers were 10.234 percentage points more likely to receive a discouraging reply. Unnamed volunteers were 1.339 percentage points more likely to receive a discouraging reply. Only the male-named effect was statistically significant $(p < .05)$.

Acknowledgements

We want to thank David Broockman, Joshua Kalla, Gabe Lenz, Tali Mendelberg, Cecilia Mo, Markus Prior, Laura Stoker, and participants at UC Berkeley's Behavior Workshop, the UC Davis American Politics Workshop, and the Nuffield College Prize Postdoctoral Research Fellow Workshop for their helpful comments and suggestions.